

More symmetric polynomials related to p -norms.

Ivo Klemeš

*Department of Mathematics and Statistics, 805 Sherbrooke Street West,
McGill University, Montréal, Québec, H3A 0B9, Canada.*

Email: klemes@math.mcgill.ca

Abstract. It is known that the elementary symmetric polynomials $e_k(x)$ have the property that if $x, y \in [0, \infty)^n$ and $e_k(x) \leq e_k(y)$ for all k , then $\|x\|_p \leq \|y\|_p$ for all real $0 \leq p \leq 1$, and moreover $\|x\|_p \geq \|y\|_p$ for $1 \leq p \leq 2$ provided $\|x\|_1 = \|y\|_1$. Previously the author proved this kind of property for $p > 2$, for certain polynomials $F_{k,r}(x)$ which generalize the $e_k(x)$. In this paper we give two additional generalizations of this type, involving two other families of polynomials. When x consists of the eigenvalues of a matrix A , we give a formula for the polynomials in terms of the entries of A , generalizing sums of principal $k \times k$ subdeterminants.

A.M.S. Mathematics Subject Classifications: 47A30 (05E05, 15A15).

Key words: inequality; p -norm; symmetric polynomial; determinant; matrix function.

Date: 17 February 2013.

§1. Introduction

Let $P_r(s) = 1 + \frac{s^1}{1!} + \cdots + \frac{s^r}{r!}$, the r th Taylor polynomial of the exponential function e^s . Let $F_{k,r}(x)$ denote the coefficient of t^k in the product

$$\prod_{i=1}^n P_r(x_i t) = \prod_{i=1}^n \left(1 + \frac{(x_i t)^1}{1!} + \cdots + \frac{(x_i t)^r}{r!} \right) = 1 + \sum_{k=1}^{nr} F_{k,r}(x) t^k, \quad (1)$$

where $x := (x_1, \dots, x_n)$ for some n . For example, when $r = 1$ we have the elementary symmetric polynomials $e_k(x) = F_{k,1}(x)$ in n variables. In [3] it was shown that the $F_{k,r}(x)$ can be used to obtain inequalities for the p -norms $\|x\|_p := (\frac{1}{n} \sum_{i=1}^n x_i^p)^{1/p}$ in certain intervals of the real number p in the following sense:

Theorem A. *Let $x, y \in [0, \infty)^n$ and fix an integer $r \geq 1$. Suppose that*

$$F_{k,r}(x) \leq F_{k,r}(y) \quad (2)$$

for all integers k in the interval $r \leq k \leq nr$. Then

$$\|x\|_p \leq \|y\|_p \quad \text{whenever} \quad 0 \leq p \leq 1. \quad (3)$$

If also $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$, then

$$\|x\|_p \geq \|y\|_p \quad \text{whenever} \quad 1 \leq p \leq r+1. \quad (4)$$

(By continuity in p , the 0-norm is defined to be the geometric mean; $\|x\|_0 := (\prod x_i)^{1/n}$.) The $F_{k,r}(x)$, together with Theorem A, may be viewed as one possible way to generalize the well-known case $r = 1$ of the theorem [1, Ch. 4, p. 211-212, Lemma 11.1], which only gives information in the range $p < 2$ and uses only the elementary symmetric polynomials. The purpose of this note is to give two other generalizations of the case $r = 1$ having the same kinds of conclusions as Theorem A in the range $p > 2$, but using two new families of symmetric polynomials, different from the above $F_{k,r}(x)$. One reason for seeking such results in the range $p > 2$ is a certain open problem on the p -norms of the eigenvalues $\{x_i\}$ of a matrix $A = QQ^*$ where Q is a $(0, 1)$ “interval matrix”. For this motivation we refer the reader to [2, Theorem 1.2] and [3, Example 1]. In this connection, one additional feature of the new polynomials is that they obey certain identities in terms of the power sum polynomials $p_m(x) = \sum_i x_i^m$ ($m = 1, 2, \dots$), similar to the Newton and “cycle index” identities for elementary symmetric polynomials. When the x_i are the eigenvalues of a matrix A these identities lead, via the Kronecker power $A^{\otimes k}$, to certain expressions for the new polynomials in terms of the entries of A . These expressions generalize the formula for $e_k(x)$ as the sum of principal $k \times k$ subdeterminants of A (see (32) to (35) in §6).

§2. Background to Theorem A.

We begin with a review of the proof of (4) in the basic case $r = 1$ in Theorem A. There is an integral formula (Mellin transform) for the power a^p of a positive real number a : For any “suitable” function $\psi(t)$, it is easily seen that

$$a^p = \frac{1}{C_p(\psi)} \int_0^\infty \psi(at) t^{-p} \frac{dt}{t}, \quad \text{where} \quad C_p(\psi) = \int_0^\infty \psi(t) t^{-p} \frac{dt}{t}. \quad (5)$$

For ψ to be “suitable”, we mean that the above improper integral $C_p(\psi)$ should converge and be nonzero. For example, with $\psi(t) = t - \log(1+t) \geq 0$, the integrals converge for $1 < p < 2$, and we have $C_p(\psi) > 0$. The restriction $p < 2$ is due to the requirement that the integrals (5) converge as $t \rightarrow 0^+$; one sees that $t - \log(1+t)$ decays like t^2 . Similarly, as $t \rightarrow \infty$, the $t - \log(1+t)$ grows like t^1 , so that the restriction $p > 1$ is needed. By applying the formula with $a = x_i$ and $a = y_i$ and summing over i , one sees that for the case $r = 1$ of (4) it is sufficient to assume

$$\sum_i \left(x_i t - \log(1 + x_i t) \right) \geq \sum_i \left(y_i t - \log(1 + y_i t) \right)$$

for all $t > 0$. For the latter, it is in turn sufficient to have the hypotheses in Theorem A; that $\sum x_i = \sum y_i$ and that each coefficient in $\prod_i (1 + x_i t)$ increases when x is replaced by y , which is exactly the hypothesis that

$F_{k,1}(x) \leq F_{k,1}(y)$ where the $F_{k,1}$ are the elementary symmetric polynomials.

Next, suppose that we want a formula for a^p valid for some $p > 2$. We could attempt to replace $t - \log(1+t)$ by a function which decays faster as $t \rightarrow 0^+$. For example, the new function $\psi(t) = t - \log(1 + t + \frac{t^2}{2!})$ can be seen to decay like t^3 , and thus yields a formula for a^p in the range $1 < p < 3$. Extending this pattern, one sees that for any positive integer r the function $\psi(t) = t - \log(1 + t + \frac{t^2}{2!} + \cdots + \frac{t^r}{r!})$ is positive and decays like t^{r+1} , and thus gives a^p in the range $1 < p < r+1$. Unravelling the required inequalities, to obtain the result (4) it is clearly sufficient to have $\sum x_i = \sum y_i$ and the polynomial inequalities $F_{k,r}(x) \leq F_{k,r}(y)$, as in Theorem A.

There are other ways to modify the function $\psi(t) = t - \log(1+t)$ to make it decay faster than t^2 as $t \rightarrow 0^+$ (while preserving some other useful aspects of the above proof). We now look at two other such modifications, thereby obtaining, after some manipulation, two new families of symmetric polynomials which can be used instead of the $F_{k,r}(x)$ in an analogous manner.

§3. The Polynomials $G_{k,r}(x)$.

Here the idea will be that instead of modifying $\log(1+t)$ from the “inside”

as was done above to obtain Theorem A, we now modify it from the “outside” and see what is obtained. Thus, we will subtract the Taylor polynomial of $\log(1+t)$ of some given degree r .

Lemma 1 *Fix an integer $r \geq 0$ and let $Q_r(t) := t - \frac{1}{2}t^2 + \cdots + (-1)^{r-1}\frac{1}{r}t^r$, the r th Taylor polynomial of $\log(1+t)$. (For $r = 0$ define $Q_0(t) := 0$.) Define*

$$\psi_r(t) := (-1)^r \left(\log(1+t) - Q_r(t) \right). \quad (6)$$

Then $\psi_r(t) > 0$ when $t > 0$, and $\psi_r(t) = \mathcal{O}(t^{r+1})$ as $t \rightarrow 0^+$. Also, for $r \geq 1$ we have $\psi_r(t) = \mathcal{O}(t^r)$ as $t \rightarrow \infty$, and for all $\epsilon > 0$ we have $\psi_0(t) = \mathcal{O}(t^\epsilon)$ as $t \rightarrow \infty$.

Proof. The proof is a standard exercise in calculus: Differentiating and then using the formula for the sum of a geometric series we obtain

$$\begin{aligned} \frac{d}{dt}\psi_r(t) &= (-1)^r \left(\frac{1}{1+t} - 1 + t - t^2 + \cdots + (-1)^{r-1}t^{r-1} \right) \\ &= (-1)^r \frac{(-t)^r}{1+t} = \frac{t^r}{1+t}. \end{aligned}$$

This shows that $\psi'_r(t) > 0$ for $t > 0$, and $\psi'_r(t) = \mathcal{O}(t^r)$ as $t \rightarrow 0^+$. Since $\psi_r(0) = 0$, the Mean Value Theorem implies that $\psi_r(t) > 0$ when $t > 0$ and

that $\psi_r(t) = \mathcal{O}(t^{r+1})$ as $t \rightarrow 0^+$. Finally, the assertions for the case $t \rightarrow \infty$ follow from the fact that the polynomial $Q_r(t)$ is of degree r , and from the growth properties of the logarithm. ■

It follows that a^p can be represented by (5) using $\psi = \psi_r$ whenever p is in the interval $r < p < r + 1$, where the corresponding constant $C_p(\psi_r)$ is positive. Hence, we deduce:

Lemma 2 *Let $x, y \in [0, \infty)^n$ and fix an integer $r \geq 0$. If*

$$\sum_{i=1}^n \psi_r(x_i t) \geq \sum_{i=1}^n \psi_r(y_i t) \tag{7}$$

for all $t > 0$, then $\|x\|_p \geq \|y\|_p$ for all p in the interval $r \leq p \leq r + 1$.

Assume now the additional hypothesis $\sum_i x_i = \sum_i y_i$. Fixing the integer $r \geq 0$, our next goal is to find a set of polynomial inequalities of the form $G_{k,r}(x) \geq G_{k,r}(y)$, or perhaps the form $G_{k,r}(x) \leq G_{k,r}(y)$, which would imply (7). Moreover, let us agree that we want these polynomials $G_{k,r}(x)$ to have positive coefficients. Before stating our result for general r , let us explain what it is for the cases $r = 0, 1, 2, 3$ in turn.

For $r = 0$, we have $\sum_i \psi_0(x_i t) = \sum_i \log(1 + x_i t)$. But $\log(1 + x_i t)$ does not have a Taylor series converging for all values of the variable t , so we cannot use the coefficient of t^k as our choice of polynomial $G_{k,0}(x)$. A simple remedy

is to exponentiate, obtaining $\prod_i(1 + x_it)$, and then use the coefficient of t^k to define the $G_{k,0}(x)$. Then, the hypothesis $G_{k,0}(x) \geq G_{k,0}(y)$ clearly implies (7) with $r = 0$. (Here we did not need the assumption $\sum_i x_i = \sum_i y_i$.)

For $r = 1$, we have $\sum_i \psi_1(x_it) = -\sum_i \left(\log(1 + x_it) - x_it \right)$. In this case, to get an entire function with positive coefficients, we first negate this, add $(\sum x_i)t$ and then exponentiate, obtaining $\prod(1 + x_it)$. We denote the coefficient of t^k by $G_{k,1}(x)$, which happens to be the same as $G_{k,0}(x)$. Note that to obtain inequality (7) for $r = 1$, we now need the hypothesis to be $G_{k,1}(x) \leq G_{k,1}(y)$, because of the negation performed at the beginning.

For $r = 2$, we have $\sum_i \psi_2(x_it) = \sum_i \left(\log(1 + x_it) - x_it + \frac{1}{2}x_i^2t^2 \right)$. To end up with positive coefficients after exponentiating, we again decide to get rid of the negative $\sum -x_it$ terms by adding on the term $(\sum x_i)t$. Hence we define $G_{k,2}(x)$ to be the coefficient of t^k in the generating function

$$\left(\prod_i (1 + x_it) \right) \exp \left(\frac{1}{2} \sum_i x_i^2 t^2 \right),$$

which we note is entire in t , whence its Taylor series converges to its value for every fixed x . Thus, the conditions $G_{k,2}(x) \geq G_{k,2}(y)$ and $\sum x_i = \sum y_i$ imply (7) with $r = 2$.

For $r = 3$, we have

$$\sum_i \psi_3(x_i t) = - \sum_i \left(\log(1 + x_i t) - x_i t + \frac{1}{2} x_i^2 t^2 - \frac{1}{3} x_i^3 t^3 \right).$$

As for $r = 1$, we first remove the leading $-$ sign. Then add on the new terms $(\sum x_i)t + \frac{1}{3}(\sum x_i)^3 t^3$ to get rid of negative coefficients. (Note that the polynomial $(\sum x_i)^3 - \sum x_i^3$ has only positive coefficients.) Therefore we let $G_{k,3}(x)$ be the coefficient of t^k in the generating function

$$\left(\prod_i (1 + x_i t) \right) \exp \left(\frac{1}{2} \sum_i x_i^2 t^2 + \frac{1}{3} \left(\left(\sum_i x_i \right)^3 - \sum_i x_i^3 \right) t^3 \right).$$

Clearly, the conditions $G_{k,3}(x) \leq G_{k,3}(y)$ and $\sum x_i = \sum y_i$ imply (7) with $r = 3$.

We continue this pattern for the general case of $r \geq 0$: In the expression

$$(-1)^r \sum_i \psi_r(x_i t) = \sum_i \left(\log(1 + x_i t) - Q_r(x_i t) \right),$$

the terms in $\sum_i -Q_r(x_i t)$ with negative coefficients are precisely those of the form $-\frac{1}{m} \sum_i x_i^m t^m$ for odd $m \leq r$. So for each such odd m we can add the extra term $+\frac{1}{m} \left(\sum_i x_i \right)^m t^m$, to make all coefficients positive. If $\sum_i x_i = \sum_i y_i$, then these extra terms are the same with y_i as with x_i . Hence, we can use the exponential of the resulting expression as a generating function to define polynomials $G_{k,r}(x)$ having positive coefficients and the other desired properties.

We now re-state the latter construction of the $G_{k,r}(x)$ using more formal notation. For notational efficiency, we observe that the odd power terms of a polynomial $f(t)$ may be written as $\frac{1}{2}(f(t) - f(-t)) =: f^-(t)$. Thus

$$Q_r^-(t) = \sum_{\text{odd } m \leq r} \frac{1}{m} t^m.$$

Define the generating function $g_r(x, t)$ by

$$g_r(x, t) := \left(\prod_{i=1}^n (1 + x_i t) \right) \exp \left(- \sum_i Q_r(x_i t) + Q_r^-\left(\sum_i x_i t\right) \right), \quad (8)$$

where we recall that $Q_r(s) := s - \frac{1}{2}s^2 + \dots + (-1)^{r-1} \frac{1}{r} s^r$, the r th Taylor polynomial of $\log(1+s)$. It is clear from (8) that $g_r(x, t)$ is an entire function of t for every fixed x .

Definition 1 *For any integers $r \geq 0, k \geq 1$, the polynomial $G_{k,r}(x)$ is the coefficient of t^k in the power series expansion of $g_r(x, t)$.*

The preceding discussion has shown that the $G_{k,r}(x)$ are symmetric polynomials with positive coefficients, and have the following property:

Theorem 1 *Let $x, y \in [0, \infty)^n$ and fix an integer $r \geq 0$. Suppose that $\sum_i x_i = \sum_i y_i$ and that for all positive integers k ,*

$$(-1)^r \left(G_{k,r}(x) - G_{k,r}(y) \right) \geq 0.$$

Then $\|x\|_p \geq \|y\|_p$ for all real p in the interval $r \leq p \leq r+1$.

In a later section (§5) we will express $G_{k,r}(x)$ in terms of the power sums $p_m = p_m(x) = \sum_{i=1}^n x_i^m$, $m = 1, 2, \dots$. To that end, we note the following alternative expression for the generating function $g_r(x, t)$ in the sense of formal power series, which is easily checked by taking logarithms in (8) and expanding each $\log(1 + x_i t)$ in powers of t :

$$g_r(x, t) = \exp \left(\sum_{m \geq 1} (-1)^{m-1} \alpha_m t^m / m \right) \quad (9)$$

where $(\alpha_m)_{m=1}^\infty$ is the “modified” sequence of power sums given by:

$$\alpha_m = \begin{cases} p_1^m & , \text{ for } m \text{ odd, } 1 \leq m \leq r, \\ 0 & , \text{ for } m \text{ even, } 1 \leq m \leq r, \\ p_m & , \text{ for } m \geq r + 1. \end{cases} \quad (10)$$

Equivalently, if r is odd, $(\alpha_m) = (p_1, 0, p_1^3, 0, \dots, 0, p_1^r, p_{r+1}, p_{r+2}, \dots)$, and if r is even, $(\alpha_m) = (p_1, 0, p_1^3, 0, \dots, p_1^{r-1}, 0, p_{r+1}, p_{r+2}, \dots)$. Note that the special cases $r = 0$ and $r = 1$ give the “full” sequence of power sums $(\alpha_m = p_m, \forall m \geq 1)$. Then $g_r(x, t)$ is of course the familiar generating function of the elementary symmetric polynomials $e_k(x)$.

In addition to the fact that $G_{k,r}(x)$ has positive coefficients, it may be interesting to note that $\pm G_{k,r}(x)$ is *Schur convex* in the sense of majorization theory [4]. (The $F_{k,r}(x)$ in our introduction also have such a property [3].) Specifically, for $r = 0$ and all odd $r \geq 1$, $-G_{k,r}(x)$ is Schur convex for all k (equivalently, $G_{k,r}(x)$ is *Schur concave*), and for all even $r \geq 2$, $G_{k,r}(x)$ is

Schur convex for all k . As discussed above, for both $r = 0, 1$ the $G_{k,r}(x)$ are just the elementary symmetric polynomials, which are well-known examples of Schur concave functions. The standard proof in the latter case consists in verifying the Schur-Ostrowski criterion, which is what we will also do below to prove our assertion for all $r \geq 1$. In fact, we prove the following slightly stronger statement involving the generating function $g_r(x, t)$.

Theorem 2 *Fix any integer $r \geq 1$ and indices $i \neq j$. Then*

$$(-1)^r \left(\frac{\partial g_r}{\partial x_i} - \frac{\partial g_r}{\partial x_j} \right) = (x_i - x_j) \gamma(x, t) \quad (11)$$

for some function $\gamma(x, t)$ (depending on r, i, j) having positive coefficients when expanded as a power series in all of its variables x_1, \dots, x_n, t . In particular, the coefficient of each t^k in $\gamma(x, t)$ is a polynomial in x_1, \dots, x_n having positive coefficients, and thus each $(-1)^r G_{k,r}(x)$ is Schur convex.

Proof. To work out the left-hand side of (11), we may let $i = 1$ and $j = 2$ by symmetry. Since $\log g_r = \sum_i \log(1 + x_i t) - \sum_i Q_r(x_i t) + Q_r^-(\sum_i x_i t)$, we obtain

$$\begin{aligned} \frac{1}{g_r} \frac{\partial g_r}{\partial x_1} &= \frac{t}{1 + x_1 t} - t Q_r'(x_1 t) + t (Q_r^-)'(\sum_i x_i t) \\ &= t (-1)^r \frac{(x_1 t)^r}{1 + x_1 t} + t (Q_r^-)'(\sum_i x_i t) \end{aligned}$$

where we have used the geometric series formula $\frac{1}{1+s} - Q'_r(s) = (-1)^r \frac{s^r}{1+s}$, as seen in the proof of Lemma 1. Similarly for x_2 . Hence

$$\begin{aligned} \frac{1}{g_r} \left(\frac{\partial g_r}{\partial x_1} - \frac{\partial g_r}{\partial x_2} \right) &= t^{r+1} (-1)^r \left(\frac{x_1^r}{1+x_1 t} - \frac{x_2^r}{1+x_2 t} \right) \\ &= t^{r+1} (-1)^r \frac{(x_1^r - x_2^r) + x_1 x_2 t (x_1^{r-1} - x_2^{r-1})}{(1+x_1 t)(1+x_2 t)}. \end{aligned}$$

Multiplying this by $(-1)^r g_r / (x_1 - x_2)$, it is now clear that $\gamma(x, t)$ has a power series with positive coefficients in all variables as claimed: The denominator $(1+x_1 t)(1+x_2 t)$ will be cancelled by the product $\prod_i (1+x_i t)$ in (8), each of the two terms of type $(x_1^m - x_2^m)/(x_1 - x_2)$ simplifies to a sum with positive coefficients, and the remaining factor $\exp \left(-\sum_i Q_r(x_i t) + Q_r^-(\sum_i x_i t) \right)$ in g_r is also a power series with positive coefficients in all variables (see (8)). ■

§4. The Polynomials $H_{k,r}(x)$.

In this variant of our topic, we modify the expression $t - \log(1+t)$, discussed in §2, by simply replacing t by t^r where $r = 2, 3, \dots$

Lemma 3 *Fix any integer $r \geq 1$ and define*

$$\phi_r(t) := t^r - \log(1+t^r). \tag{12}$$

Then $\phi_r(t) > 0$ when $t > 0$, $\phi_r(t) = \mathcal{O}(t^{2r})$ as $t \rightarrow 0^+$, and $\phi_r(t) = \mathcal{O}(t^r)$ as $t \rightarrow \infty$.

The lemma follows from the case $r = 1$ of Lemma 1 by substituting t^r for t . It follows that a^p can be represented by (5) using $\psi = \phi_r$ whenever p is in the interval $r < p < 2r$, where the corresponding constant $C_p(\phi_r)$ is again positive. Hence, we deduce:

Lemma 4 *Let $x, y \in [0, \infty)^n$ and fix an integer $r \geq 1$. If*

$$\sum_{i=1}^n \phi_r(x_i t) \geq \sum_{i=1}^n \phi_r(y_i t) \quad (13)$$

for all $t > 0$, then $\|x\|_p \geq \|y\|_p$ for all p in the interval $r \leq p \leq 2r$.

Next, assume the additional hypothesis $\sum_i x_i = \sum_i y_i$. Fixing the integer $r \geq 1$, we once again wish to obtain (13) as a consequence of some stronger family of symmetric polynomial inequalities, say of the form $H_{k,r}(x) \leq H_{k,r}(y)$ for some family H having positive coefficients. By the idea already seen in §3, the following generating function $h_r(x, t)$ seems natural for this purpose:

$$h_r(x, t) := \left(\prod_i (1 + x_i^r t) \right) \exp \left(\left(\left(\sum_i x_i \right)^r - \sum_i x_i^r \right) t \right). \quad (14)$$

Note that here we have in effect replaced the t^r by t , for the sake of simplicity of our generating function. As before, it is clear that $h_r(x, t)$ is entire in t for every x .

Definition 2 For any integers $r \geq 1, k \geq 1$, the polynomial $H_{k,r}(x)$ is the coefficient of t^k in the power series expansion of $h_r(x, t)$.

It is easily seen from the preceding discussion that the $H_{k,r}(x)$ are symmetric polynomials with positive coefficients, and have the following property:

Theorem 3 Let $x, y \in [0, \infty)^n$ and fix an integer $r \geq 1$. Suppose that $\sum_i x_i = \sum_i y_i$ and that for all positive integers k ,

$$H_{k,r}(x) \leq H_{k,r}(y).$$

Then $\|x\|_p \geq \|y\|_p$ for all real p in the interval $r \leq p \leq 2r$.

As in the previous section §3, for $r \geq 1$ we may express the generating function (14) in terms of the power sums p_m using formal power series:

$$h_r(x, t) = \exp \left(\sum_{m \geq 1} (-1)^{m-1} \alpha_m t^m / m \right) \quad (15)$$

where $(\alpha_m)_{m=1}^\infty$ is a modified sequence of power sums given by:

$$\alpha_m = \begin{cases} p_1^r & , \text{ for } m = 1, \\ p_{mr} & , \text{ for } m \geq 2. \end{cases} \quad (16)$$

Equivalently $(\alpha_m) = (p_1^r, p_{2r}, p_{3r}, p_{4r}, \dots)$.

Also as in §3, the polynomials $H_{k,r}(x)$ turn out to have a Schur convexity property (Schur concavity, in fact), “inherited” from their generating function:

Theorem 4 Fix any integer $r \geq 1$ and indices $i \neq j$. Then

$$\frac{\partial h_r}{\partial x_i} - \frac{\partial h_r}{\partial x_j} = -(x_i - x_j)\delta(x, t) \quad (17)$$

for some function $\delta(x, t)$ (depending on r, i, j) having positive coefficients when expanded as a power series in all of its variables x_1, \dots, x_n, t . In particular, the coefficient of each t^k in $\delta(x, t)$ is a polynomial in x_1, \dots, x_n having positive coefficients, and thus each $H_{k,r}(x)$ is Schur concave.

Proof. As in the proof of Theorem 2, we may let $i = 1$ and $j = 2$ by symmetry. Since $\log h_r = \sum_i \log(1 + x_i^r t) - \sum_i x_i^r t + (\sum_i x_i)^r t$, we obtain

$$\frac{1}{h_r} \frac{\partial h_r}{\partial x_1} = \frac{rx_1^{r-1}t}{1 + x_1^r t} - rx_1^{r-1}t + r\left(\sum_i x_i\right)^{r-1}t = \frac{-rx_1^{2r-1}t^2}{1 + x_1^r t} + r\left(\sum_i x_i\right)^{r-1}t.$$

Similarly for x_2 . Hence

$$\begin{aligned} \frac{1}{h_r} \left(\frac{\partial h_r}{\partial x_1} - \frac{\partial h_r}{\partial x_2} \right) &= -rt^2 \left(\frac{x_1^{2r-1}}{1 + x_1^r t} - \frac{x_2^{2r-1}}{1 + x_2^r t} \right) \\ &= -rt^2 \frac{(x_1^{2r-1} - x_2^{2r-1}) + x_1^r x_2^r t (x_1^{r-1} - x_2^{r-1})}{(1 + x_1^r t)(1 + x_2^r t)}. \end{aligned}$$

Multiplying this by $h_r/(x_1 - x_2)$, it is clear that $\delta(x, t)$ has a power series with positive coefficients in all variables as claimed, in view of (14). ■

§5. Expressions in terms of power sums.

Our next aim is to express $G_{k,r}(x)$ and $H_{k,r}(x)$ in terms of the power sums $p_m(x) = \sum_{i=1}^n x_i^m$ where m is a positive integer. This follows immediately from (9) and (15) by well-known formulas for the exponential of a power series and the (signed) cycle index polynomials Z_k , which we now recall for the reader's convenience in the form of Lemma 5 below. Let $\alpha_1, \alpha_2, \dots$ be any formal commuting variables or “indeterminates”. If $\lambda = (\lambda_j) = (\lambda_1 \leq \lambda_2 \leq \dots)$ is a partition of k we define $\alpha_\lambda = \prod_j \alpha_{\lambda_j}$. Define the sign of λ by $\text{sgn}(\lambda) = \prod_j (-1)^{\lambda_j-1}$ ($= \beta_\lambda$ for the special sequence $\beta_n = (-1)^{n-1}$). Let $n_m(\lambda) \geq 0$ denote the number of j 's such that $\lambda_j = m$, and let \mathcal{P}_k denote the set of all partitions λ of k . If σ is an element of S_k , the group of permutations of $\{1, \dots, k\}$, and if $(\lambda_j) =: \lambda(\sigma)$ are the lengths of the cycles in the disjoint cycle decomposition of σ , define $\alpha_\sigma = \alpha_{\lambda(\sigma)}$. Also, define $n_m(\sigma) = n_m(\lambda(\sigma))$, and $\text{sgn}(\sigma) = \text{sgn}(\lambda(\sigma))$, the usual sign of a permutation.

Lemma 5 [“Exponential Formula”] *If $\alpha = (\alpha_m)_{m=1}^\infty$ is any sequence of indeterminates, then the following identity holds in the sense of formal power series in t .*

$$\exp \left(\sum_{m \geq 1} (-1)^{m-1} \alpha_m t^m / m \right) = 1 + \sum_{k \geq 1} Z_k(\alpha_1, \dots, \alpha_k) t^k, \quad (18)$$

where the polynomials Z_k are given by:

$$Z_k(\alpha_1, \dots, \alpha_k) = \sum_{\lambda \in \mathcal{P}_k} \text{sgn}(\lambda) \prod_{m=1}^k \frac{(\alpha_m/m)^{n_m(\lambda)}}{n_m(\lambda)!} \quad (19)$$

$$= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_\sigma \quad (20)$$

$$= \frac{1}{k!} \begin{vmatrix} \alpha_1 & 1 & 0 & 0 & \cdot & \cdot & 0 \\ \alpha_2 & \alpha_1 & 2 & 0 & \cdot & \cdot & 0 \\ \alpha_3 & \alpha_2 & \alpha_1 & 3 & \cdot & \cdot & 0 \\ \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & (k-1) \\ \alpha_k & \alpha_{k-1} & \cdot & \cdot & \cdot & \cdot & \alpha_1 \end{vmatrix}. \quad (21)$$

A convenient reference for this result and further background is [7, Theorem 1.3.3]. The determinant (21) is equivalent to a recursion (or “Newton identities”) given in [8, Proposition 5.1.7]. (One can also recover (20) and (21) from [5, Eqn. (8.30) and (8.31)], which is the special case $\alpha_m = p_m, \forall m \geq 1$ and $Z_k = e_k, \forall k \geq 1$. This implies the case of arbitrary α_m by the algebraic independence of $p_1(x), \dots, p_k(x)$ for $n \geq k$.)

Corollary 1 *Fix an integer $r \geq 0$. Then for all $k \geq 1$,*

$$G_{k,r} = Z_k(\alpha_1, \dots, \alpha_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_\sigma \quad (22)$$

where $(\alpha_m)_{m=1}^\infty$ is the “modified” sequence of power sums given by (10). Let the subset $S(k, r) \subset S_k$ consist of those permutations σ having no cycles of

even length $\leq r$ (equivalently, $n_m(\sigma) = 0$ for all even $m \leq r$). Then (22)

becomes

$$G_{k,r} = \frac{1}{k!} \sum_{\sigma \in S(k,r)} \text{sgn}(\sigma) p_1^{\sum_{j \leq r} j n_j(\sigma)} \prod_{j=r+1}^k p_j^{n_j(\sigma)} . \quad (23)$$

Corollary 2 Fix an integer $r \geq 1$. Then for all $k \geq 1$,

$$H_{k,r} = Z_k(\alpha_1, \dots, \alpha_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_\sigma \quad (24)$$

where $(\alpha_m)_{m=1}^\infty$ is the “modified” sequence of power sums given by (16). Substituting (16), this becomes

$$H_{k,r} = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) p_1^{r n_1(\sigma)} \prod_{j=2}^k p_j^{n_j(\sigma)} . \quad (25)$$

The results (23) and (25) may also be expressed as follows. For each $\sigma \in S(k, r)$, let σ_r denote the product of all cycles of length $\geq r+1$ in the disjoint cycle factorization of σ . By convention, an empty product is taken to mean the identity of the group S_k . (In terms of the mappings, $\sigma_r = \sigma$ on the σ -orbits of length $\geq r+1$, and $\sigma_r =$ the identity map on the σ -orbits of length $\leq r$.) Note that $\text{sgn}(\sigma_r) = \text{sgn}(\sigma)$ for $\sigma \in S(k, r)$. Then (23) may be written as

$$G_{k,r} = \frac{1}{k!} \sum_{\sigma \in S(k,r)} \text{sgn}(\sigma_r) p_{\sigma_r} . \quad (26)$$

Regarding (25), we may alternatively regard $H_{k,r}$ as the coefficient of $t^{(kr)}$ in the generating function

$$h_r(x, t^r) = \exp \left(\sum_{j \geq 1} \alpha_j (-1)^{j-1} t^{jr} / j \right).$$

Applying the Exponential Formula (Lemma 5) to the latter and doing the arithmetic leads to

$$H_{k,r} = \frac{(-1)^{k(r-1)}}{(kr)!} \sum_{\sigma \in T(k,r)} \text{sgn}(\sigma) r^{L(\sigma)} p_{\sigma_r}, \quad (27)$$

where $T(k, r) \subset S_{kr}$ consists of those permutations σ of $\{1, \dots, kr\}$ whose disjoint cycles all have lengths divisible by r , $L(\sigma)$ denotes the number of disjoint cycles, and σ_r is again the product of all cycles of length $> r$ in the disjoint cycle factorization of σ (i.e. each r -cycle of $\sigma \in T(k, r)$ is replaced by r 1-cycles).

Remark. Fix the integer $r \geq 1$. In the ring of symmetric functions (in infinitely many variables), we may define an algebraic homomorphism ϕ by arbitrarily defining the image of each power sum p_m , since these constitute a basis. Hence we may define a homomorphism by $\phi(p_m) = \alpha_m$ for all $m \geq 1$, where α_m is given by (10). Then $\phi(e_k) = G_{k,r}$ for all $k \geq 1$, by Corollary 1. (Since $\phi(e_k) = \phi(Z_k(p_1, \dots, p_k)) = Z_k(\phi(p_1), \dots, \phi(p_k)) = Z_k(\alpha_1, \dots, \alpha_k) = G_{k,r}$.) Similarly, by Corollary 2, the $H_{k,r}$ are the images of the e_k under the

homomorphism ψ defined by $p_m \mapsto \alpha_m$ where now α_m means (16). We digress to mention the following question regarding ψ . Let \mathcal{S} denote the set of all linear combinations, with positive real coefficients, of all Schur functions (see [6, §4.4]). Is $\psi(\mathcal{S}) \subset \mathcal{S}$? (In particular, is each $H_{k,r}$ a linear combination, with positive real coefficients, of Schur functions?)

§6. Expressions in terms of matrix entries.

Let $A = [a_{ij}]$ be a complex $n \times n$ matrix, and let $(x_1, \dots, x_n) = x$ be its eigenvalues. In this section we briefly consider the question of how to express $G_{k,r}(x)$ and $H_{k,r}(x)$ as polynomials in the entries a_{ij} of A . This is possible of course for *any* symmetric polynomial $F(x)$, since by a fundamental result $F(x)$ may first be written as a polynomial in the power sums $p_k(x)$ (or, if we prefer, in the elementary symmetric polynomials $e_k(x)$), and these in turn have well-known polynomial expressions in terms of the entries a_{ij} ; $p_k(x) = \text{Trace}(A^k) := \langle A^k \rangle$ and $e_k(x) =$ sum of all principal $k \times k$ subdeterminants of A . We can therefore already give one answer quite explicitly using the polynomials $Z_k(\alpha)$ from Corollaries 1 or 2 of the previous section, replacing each occurrence of a p_m by the trace $\langle A^m \rangle$. For example we thus obtain:

$$G_{5,3}(x) = Z_5(p_1, 0, p_1^3, p_4, p_5) = Z_5(\langle A \rangle, 0, \langle A \rangle^3, \langle A^4 \rangle, \langle A^5 \rangle)$$

$$\begin{aligned}
&= \frac{1}{5!} \sum_{\sigma \in S(5,3)} \text{sgn}(\sigma) \langle A \rangle^{n_1(\sigma)+3n_3(\sigma)} \langle A^4 \rangle^{n_4(\sigma)} \langle A^5 \rangle^{n_5(\sigma)} \\
&= \frac{7}{40} \langle A \rangle^5 - \frac{1}{4} \langle A \rangle \langle A^4 \rangle + \frac{1}{5} \langle A^5 \rangle .
\end{aligned}$$

However, we will now present another kind of expression which uses more explicitly the individual monomials (products) of the entries a_{ij} of A . For example, $G_{k,1}(x) = e_k(x)$ is on the one hand given by $Z_k(\langle A \rangle, \langle A^2 \rangle, \dots, \langle A^k \rangle)$, but on the other hand $e_k(x)$ is also equal to the sum of all principal $k \times k$ subdeterminants of A , which can immediately be written in terms of products of a_{ij} using the familiar expansion of a determinant as a sum over permutations. We thus aim to generalize this latter type of expansion for any $G_{k,r}$ or $H_{k,r}$, and will do so essentially by imitating what happens in the special case of $e_k = G_{k,1}$. We need some of the machinery of representation theory. Following the sketch given in [8, Ch. 7, Appendix 2, pp. 444-445], or the more detailed account [5, Ch. 6], identify A with the linear map $A : V \rightarrow V$ of the vector space $V = \mathbb{C}^n$ (of column vectors) as usual, and consider the k -fold tensor product $V^{\otimes k} = V \otimes \cdots \otimes V$. There is an action of A which we denote by $A^{\otimes k}$ (the “Kronecker power”) from $V^{\otimes k}$ to itself, characterized by $A^{\otimes k}(v_1 \otimes \cdots \otimes v_k) = (Av_1) \otimes \cdots \otimes (Av_k)$ for all elements of the form $v_1 \otimes \cdots \otimes v_k$. Also, there is an action for any $\sigma \in S_k$,

characterized by $\sigma(v_1 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$. The two actions commute; $\sigma A^{\otimes k} = A^{\otimes k} \sigma$ on $V^{\otimes k}$. The trace of $A^{\otimes k} \sigma$ on $V^{\otimes k}$ depends only on the conjugacy class of σ and eigenvalues x of A :

$$\text{Trace}(A^{\otimes k} \sigma) = p_\sigma(x) = \prod_j p_{\lambda_j(\sigma)}(x), \quad (28)$$

where the p_m are the power sum polynomials as before. (This can be seen by using a basis of V which makes A triangular.) By linearity, for any sequence of permutations $\sigma_s \in S_k$ and constants $c_s \in \mathbb{C}$ on some finite index set S we have

$$\text{Trace}\left(\sum_{s \in S} c_s A^{\otimes k} \sigma_s\right) = \sum_{s \in S} c_s p_{\sigma_s}(x). \quad (29)$$

On the other hand, we may also compute a trace on $V^{\otimes k}$ using the basis $\{u_{m(1)} \otimes \cdots \otimes u_{m(k)}\}$ where $\{u_j\}_{j=1}^n$ is the standard basis of $V = \mathbb{C}^n$, and $m \in \Gamma_{k,n} = \text{set of all functions } m : \{1, \dots, k\} \rightarrow \{1, \dots, n\}$. This will yield the above trace (29) directly in terms of products of entries a_{ij} of A . This computation is well-known, but we will now reproduce it here for convenience. We will essentially follow [5, Ch. 6 and 7]. For a given pair of sequences $(c_s, \sigma_s)_{s \in S} =: C$ as above, define a “matrix function” d_C on complex $k \times k$ matrices $B = [B_{i,j}]$ by

$$d_C(B) = \sum_{s \in S} c_s \prod_{j=1}^k B_{\sigma_s(j), j} \quad (30)$$

Now consider a fixed $m \in \Gamma_{k,n}$. Recalling that u_j is the j th standard basis column vector in \mathbb{C}^k , denote

$$u(m) := u_{m(1)} \otimes u_{m(2)} \otimes \cdots \otimes u_{m(k)}.$$

Denote the entries of A by $A(i, j) := a_{i,j}$ for the sake of better legibility. Let $\langle v, w_j \rangle$ denote the coefficient of w_j in v whenever v is an element of a vector space of which $\{w_j\}$ is a basis. Then for each $\sigma \in S_k$ we find that

$$\left\langle A^{\otimes k} \sigma \left(u(m) \right), u(m) \right\rangle = \prod_{j=1}^k A \left(m(\sigma(j)), m(j) \right) =: \prod_{j=1}^k A[m|m]_{\sigma(j),j}$$

where $A[m|m]$ denotes the $k \times k$ matrix whose (i, j) entry is $A(m(i), m(j))$.

We remark that $A[m|m]$ is a (principal) submatrix of A only when the function m is strictly increasing, otherwise it may be thought of as a “submatrix” which allows repetition and permutation of the original indices. Hence, for each fixed $m \in \Gamma_{k,n}$,

$$\sum_{s \in S} c_s \left\langle A^{\otimes k} \sigma_s \left(u(m) \right), u(m) \right\rangle = d_C(A[m|m]).$$

Summing over all m gives

$$\text{Trace} \left(\sum_{s \in S} c_s A^{\otimes k} \sigma_s \right) = \sum_{m \in \Gamma_{k,n}} d_C(A[m|m]).$$

Hence by (29),

$$\sum_{s \in S} c_s p_{\sigma_s}(x) = \sum_{m \in \Gamma_{k,n}} d_C(A[m|m]). \quad (31)$$

We call $C = (c_s, \sigma_s)_{s \in S}$ a *class function* on S_k if the index set $S = S_k$, $\sigma_s = s$, and c_s depends only on the conjugacy class of s . When C is a class function, the latter sum may be simplified further by noting that $d_C(A[m|m]) = d_C(A[m'|m'])$ whenever m and m' are permutations of each other, i.e. when $m' = m \circ \tau$ for some $\tau \in S_k$. Let $(N_1, \dots, N_n) =: N$ denote an n -tuple of integers $N_i \geq 0$ with $\sum_i N_i = k$, and let $m^{(N)} \in \Gamma_{k,n}$ denote the unique nondecreasing function which takes on each value i exactly N_i times. Then there are exactly $k!/(N_1! \dots N_n!)$ distinct permutations of $m^{(N)}$ in $\Gamma_{k,n}$. Hence, whenever a matrix function d_γ can be shown to be of the form $d_\gamma = d_C$ for some class function C , then we have

$$\sum_{m \in \Gamma_{k,n}} d_\gamma(A[m|m]) = \sum_{\sum_i N_i = k} d_\gamma(A[m^{(N)}|m^{(N)}]) \frac{k!}{N_1! \dots N_n!}.$$

For our cases of interest, it remains to apply (31) to the results (26) or (27). In the first case, one obtains

$$G_{k,r}(x) = \frac{1}{k!} \sum_{m \in \Gamma_{k,n}} \delta_r(A[m|m]) = \sum_{\sum_i N_i = k} \delta_r(A[m^{(N)}|m^{(N)}]) \frac{1}{N_1! \dots N_n!}, \quad (32)$$

where $\delta_r(B)$ is a “modified determinant” defined for any $k \times k$ matrix B by

$$\delta_r(B) = \sum_{\sigma \in S(k,r)} \text{sgn}(\sigma_r) \prod_{j=1}^k B_{\sigma_r(j),j}, \quad (33)$$

which is easily seen to be of the form $d_C(B)$ for some class function C on S_k .

Similarly, from (27) one obtains

$$H_{k,r}(x) = \frac{1}{(kr)!} \sum_{m \in \Gamma_{kr,n}} D_r(A[m|m]) = \sum_{\sum_i N_i = kr} D_r(A[m^{(N)}|m^{(N)}]) \frac{1}{N_1! \dots N_n!}, \quad (34)$$

where $D_r(B)$ is defined for any $kr \times kr$ matrix B as:

$$D_r(B) = (-1)^{k(r-1)} \sum_{\sigma \in T(k,r)} \text{sgn}(\sigma) r^{L(\sigma)} \prod_{j=1}^{kr} B_{\sigma_r(j),j}, \quad (35)$$

which can be checked to be $d_C(B)$ for some class function C on S_{kr} .

References

- [1] Gohberg, I. C. and Kreĭn, M. G. *Introduction to the theory of linear nonselfadjoint operators*. Translated from the Russian by A. Feinstein. Transl. of Math. Monographs, Vol. 18, American Math. Soc., Providence, R.I., 1969.
- [2] Klemeš, I. *Finite Toeplitz matrices and sharp Littlewood conjectures*. (English) Algebra i Analiz 13 (2001), no. 1, pp. 39-59; translation in St. Petersburg Math. J. 13 (2002), no. 1, 27-40.
- [3] Klemeš, I. *Symmetric polynomials and l^p inequalities for certain intervals of p* . Houston J. Math. Vol. 37, (2011), no. 1, pp. 285-295.

- [4] Marshall, A. W. and Olkin, I. *Inequalities: theory of majorization and its applications*. Academic Press, New York-London, 1979.
- [5] Merris, R. *Multilinear algebra*. Algebra, Logic and Applications, 8. Gordon and Breach Science Publishers, Amsterdam, 1997.
- [6] Sagan, B. E. *The symmetric group. Representations, combinatorial algorithms, and symmetric functions*. 2nd ed. Graduate Texts in Mathematics, 203. Springer-Verlag, New York, 2001.
- [7] Stanley, R. P. *Enumerative combinatorics*. Vol. 1, 2nd ed. Cambridge Studies in Advanced Mathematics, 49. Cambridge University Press, Cambridge, 2012.
- [8] Stanley, R. P. *Enumerative combinatorics*. Vol. 2, Cambridge Studies in Advanced Mathematics, 62. Cambridge University Press, Cambridge, 1999.